MODIS Level 2 IMAPP Atmospheric Science Processing Algorithm

IMAPP_SPA

General

The NASA Goddard Space Flight Center's (GSFC) Direct Readout Laboratory (DRL), Code 606.3 developed this wrapper software for the National Polar-orbiting Operational Environmental Satellite System (NPOESS) Preparatory Project (NPP) In-Situ Ground System (NISGS) and the International Polar Orbiter Processing Package (IPOPP).

Users must agree to all terms and conditions in the Software Usage Agreement on the DRL Web Portal before downloading this software.

Software and documentation published on the DRL Web Portal may occasionally be updated or modified. The most current versions of DRL software are available at the DRL Web Portal:

http://directreadout.sci.gsfc.nasa.gov

Questions relating to the contents or status of this software and its documentation should be addressed to the DRL via the Contact Us mechanism at the DRL Web Portal:

http://directreadout.sci.gsfc.nasa.gov/index.cfm?section=contact%20us

Algorithm Wrapper Concept

The DRL has developed an algorithm wrapper to provide a common command and execution interface to encapsulate multi-discipline, multi-mission science processing algorithms. The wrapper also provides a structured, standardized technique for packaging new or updated algorithms with minimal effort.

A Science Processing Algorithm (SPA) is defined as a wrapper and its contained algorithm. SPAs will function in a standalone, cross-platform environment to serve the needs of the broad Direct Readout community. Detailed information about SPAs and other DRL technologies is available at:

http://directreadout.sci.gsfc.nasa.gov/index.cfm?section=technology

Software Description

This DRL software package contains the MODIS IMAPP_SPA (International MODIS/AIRS Processing Package SPA). The IMAPP_SPA processes MODIS Aqua and Terra Level 1B Direct Broadcast (DB) data into four Level 2 MODIS atmospheric products: Cloudmask (MOD35); Cloudtop Properties, Cloud Phase and Cloud Optical properties (MOD06); Atmospheric Profiles (MOD07); and Aerosol (MOD04). The IMAPP_SPA functions in two modes: Standalone, or as an IPOPP plug-in.

Software Version

Version 1.1 of the DRL algorithm wrapper was used to package the SPA described in this document. The SPA uses IMAPP Version 2.1 (October 17, 2009) processing code to generate MODIS Level 2 Atmospheric products, and incorporates MODIS Destripe Direct Broadcast Software Version 1.0 (September 24, 2008).

Enhancements to this SPA include:

a) The Cloud Optical Properties software module is now included. This software module will create 19 optical property arrays at 1km resolution, including Cloud_Optical_Thickness and Cloud_Effective_Radius, and place them into the MOD06 output product.

b) The Cloud Top Properties software module has been updated to Collection 5 MODAPS Version 5.3.1. It now requires an 8-day Clear Sky Radiance Bias (CSRB) ancillary file for creation of the MOD06 product.

c) The Atmospheric Profiles algorithm now creates a Retrieved_Ozone_Profile dataset in the MOD07 product. It also retrieves Moisture Profile in units of mixing ratio (g/kg) instead of Dewpoint (°C). Corresponding coefficient, bias and detector static ancillary files have been updated.

d) The grib2 meteorological ancillary-ingest capability has been extended (the grib2 ancillary-ingest capability is limited to 64-bit Linux platforms and is presently available in Standalone Mode only).

e) Ancillary logic for optimal selection of ancillary files during real time processing and reprocessing in IPOPP Mode has been improved.

This software will execute on 64- and 32-bit computers, and has been tested with the following operating systems: Fedora 12, CentOS 5.4, Kubuntu 10.04, and SUSE 11.2.

Credits

The IMAPP and MODIS Destripe Direct Broadcast software packages were provided to the DRL by the Space Science and Engineering Center (SSEC), University of Wisconsin-Madison.

Prerequisites

To run this package, you must have the Java Development Kit (JDK) or Java Runtime Engine (JRE) (Java 1.5 or higher) installed on your computer, and the bin directory of your Java installation in your PATH environment variable.

Program Inputs and Outputs

Inputs to the IMAPP_SPA are as follows:

a) the MODIS 1km, half km, and quarter km L1B Calibrated Geolocated Radiances Hierarchical Data Format (HDF) products;

b) the MODIS Geolocation HDF product;

c) ancillary files for ice/snow extent, sea ice concentration, sea surface temperature, ozone, clear sky radiance bias, and weather.

The IMAPP_SPA outputs the following MODIS Level 2 atmospheric products:

a) Cloudmask (MOD35);

b) Cloudtop Properties, Cloud Phase and Cloud Optical Properties (MOD06);

c) Atmospheric Profiles (MOD07);

d) Aerosol (MOD04).

Installation and Configuration

This section contains instructions for installing an SPA in a standalone configuration. SPAs may also be installed dynamically into an IPOPP framework; instructions for this type of installation are contained in the IPOPP User's Guide.

Download the IMAPP_2.1_SPA_1.1.tar.gz,
IMAPP_2.1_SPA_1.1_MOD06OD_COEFF.tar.gz and
IMAPP_2.1_SPA_1.1_testdata.tar.gz (optional) files into the same directory.

Decompress and un-archive the IMAPP_2.1_SPA_1.1.tar.gz,
IMAPP_2.1_SPA_1.1_MOD06OD_COEFF.tar.gz and
IMAPP_2.1_SPA_1.1_testdata.tar.gz (optional) files:

```
$ tar -xzf IMAPP_2.1_SPA_1.1.tar.gz
$ tar -xzf  IMAPP_2.1_SPA_1.1_MOD06OD_COEFF.tar.gz
$ tar -xzf IMAPP_2.1_SPA_1.1_testdata.tar.gz
```

This will create the following subdirectories:

SPA

IMAPP

algorithm

ancillary

wrapper

stations

testscripts

testdata

NOTES:

1. Examples supplied from this point forward assume that the SPA was installed into /home/ipopp/drl/, although in Standalone Mode SPAs may be installed into any directory.
2. The IMAPP_2.1_SPA_1.1_MOD06OD_COEFF.tar.gz contains coefficient files required by the MODIS Cloud Optical Properties software module (MOD06OD). This module produces 19 optical property arrays at 1km resolution, including Cloud_Optical_Thickness and Cloud_Effective_Radius, and places them into the MOD06 output product. In the absence of these coefficient files, the MOD06OD will not be run and the corresponding datasets in the MOD06 output product will remain unpopulated.
3. Decompressing and un-archiving the IMAPP_2.1_SPA_1.1_MOD06OD_COEFF.tar.gz file will place the MOD06OD coefficients into the /home/ipopp/drl/SPA/IMAPP/algorithm/src/cloudoptical/coefficients directory. Once these files have been extracted, please make sure they are available through the /home/ipopp/drl/SPA/IMAPP/algorithm/coeff directory (there are soft links waiting for the Forward_Libraries, Surface_Albedo and Water_Vapor_Transmittance directories). If you do not see these subdirectories, please soft-link them into the /home/ipopp/drl/SPA/IMAPP/algorithm/coeff directory, like this:

   ```
   cd /home/ipopp/drl/SPA/IMAPP/algorithm/coeff
   ln -s ../src/cloudoptical/coefficients/Forward_Libraries .
   ln -s ../src/cloudoptical/coefficients/Surface_Albedo .
   ln -s ../src/cloudoptical/coefficients/Water_Vapor_Transmittance .
   ```

Software Package Testing and Validation

The testscripts subdirectory contains test scripts that can be used to check that your current installation of the SPA is working properly, as described below. Note that the optional IMAPP_2.1_SPA_1.1_testdata.tar.gz file is required to execute these testing procedures.

Step 1: cd into the /home/ipopp/drl/SPA/IMAPP/testscripts directory

Step 2: There are three scripts inside the testscripts directory: 'run-imapp', 'run-imapp-autodownload' and 'run-imapp-grib2input'. The 'run-imapp' script uses input L1B files and ancillary files in the /home/ipopp/drl/SPA/IMAPP/testdata/input directory to create the atmospheric products. The 'run-imapp-autodownload' uses the same L1B files; however it automatically downloads the appropriate ancillary files needed to process the L1B files. The 'run-imapp-grib2input' script shows how to use grib2 meteorological ancillary as input. Run each one of them as follows:

```
$./run-imapp

$./run-imapp-autodownload

$./run-imapp-grib2input
```

A successful execution usually takes some time (between 20 and 30 minutes, depending on the speed of your computer), so if the execution seems to get stuck with an "Awaiting Completion" message, do not become impatient. If everything is working properly, the scripts will terminate with a message such as:

Output modis.cloudmask is
/home/ipopp/drl/SPA/IMAPP/testdata/output/mod35.10060152841.hdf

Output modis.cloudtop is
/home/ipopp/drl/SPA/IMAPP/testdata/output/mod06.10060152841.hdf

Output modis.aerosols is
/home/ipopp/drl/SPA/IMAPP/testdata/output/mod04.10060152841.hdf

Output modis.atmprofile is
/home/ipopp/drl/SPA/IMAPP/testdata/output/mod07.10060152841.hdf

You can cd to the output directory to verify that the science product exists. Test output product(s) are available for comparison in the testdata/output directory. These test output product(s) were generated on a 64-bit PC architecture computer running Fedora 10. The output products serve as an indicator of expected program output. Use a comparison utility (such as diff, hdiff, etc.) to compare your output product(s) to those

provided in the testdata/output directory. Locally generated files may differ slightly from the provided output files because of differences in machine architecture or operating systems.

NOTE: Datasets produced by the MODIS Cloud Optical Properties software module would remain unpopulated in the MOD06 product if the IMAPP_2.1_SPA_1.1_MOD06OD_COEFF.tar.gz is not installed. In that case the MOD06 output product generated by the testscripts would differ from the corresponding test output product.

If there is a problem and the code terminates abnormally, the problem can be identified using the log files. Log files are generated automatically within the directory used for execution. They start with stdfile* and errfile*. Please report any errors that cannot be fixed to the DRL. You can delete the log files after execution.

Program Operation

In order to run the package using your own input data, you can either use the 'run' scripts within the wrapper subdirectories, or modify the test scripts within the testscripts subdirectory.

To Use the Run Scripts

Identify the 'run' scripts: The /home/ipopp/drl/SPA/IMAPP/wrapper/IMAPP subdirectory contains an executable called 'run'. Execute 'run' to execute the IMAPP_SPA and create MODIS atmospheric products. Note that to execute 'run', you need to have java on your path.

Specify input parameters using <label value> pairs: To execute the 'run' scripts, you must supply the required input and output parameters. Input and output parameters are usually file paths or other values (e.g., platform name or scan time). Each parameter is specified on the command line by a <label value> pair. Labels are simply predefined names for parameters. Each label must be followed by its actual value. Each process has its own set of <label value> pairs that must be specified in order for it to execute. Some of these pairs are optional, meaning the process would still be able to execute even if that parameter is not supplied. There are three types of <label value> pairs that the MODIS IMAPP_SPA uses, as follows:

a) Input file label/values. These are input file paths. Values are absolute or relative paths to the corresponding input file.

b) Parameter label/values. These are parameters that need to be passed into the SPA (e.g., platform name or scan time).

c) Output file labels. These are output files that are produced by the SPA. Values are the relative/absolute paths of the files you want to generate.

The following tables contains labels, and their descriptions, required by the MODIS IMAPP_SPA.

Input File Labels	Description	Data Sources
modis.mxd021km	MODIS 1km L1B Calibrated Geolocated Radiances HDF file (MOD021KM, MYD021KM)	L1B files over the eastern US region can be found at: ftp://is.sci.gsfc.nasa.gov/gsfcdata/aqua/modis/level1 (for Aqua) ftp://is.sci.gsfc.nasa.gov/gsfcdata/terra/modis/level1 (for Terra) L1B files over a different region can be obtained from the DAACs or your local Direct Broadcast station.
modis.mxd02hkm	MODIS 500m L1B Calibrated Geolocated Radiances HDF file (MOD02HKM, MYD02HKM)	
modis.mxd02qkm	MODIS 250m L1B Calibrated Geolocated Radiances HDF file (MOD02QKM, MYD02QKM)	
modis.mxd03	MODIS Geolocation hdf file (MOD03, MYD03)	
ssmi_nise (optional)	NSIDC NISE (Near-real time Ice and Snow Extent) (1 degree, global, daily)	Current Data: ftp://is.sci.gsfc.nasa.gov/ancillary/temporal/global/nise Archived Data: ftp://is.sci.gsfc.nasa.gov/ArchivedAncillary/temporal/global/seaice
ssmi_seaice (optional)	National Centers for Environmental Prediction (NCEP) sea ice concentration (1 degree, global, daily)	Current Data: ftp://is.sci.gsfc.nasa.gov/ancillary/temporal/global/nise Archived Data: ftp://is.sci.gsfc.nasa.gov/ArchivedAncillary/temporal/global/seaice

Input File Labels	Description	Data Sources
ncep_met (optional)	NCEP Numerical Weather Prediction GRIdded Binary (GRIB) File. This can be either a Global Data Assimilation System (GDAS1, 6 hourly, 1 degree global) analysis field file or a Global Model Forecast Fields (GFS) file. GDAS and GFS files are available in both grib1 and grib2 formats.	Current Data: ftp://is.sci.gsfc.nasa.gov/ancillary/temporal/global/gdas (for GDAS) ftp://is.sci.gsfc.nasa.gov/ancillary/temporal/global/gfs (for GFS) Archived Data: ftp://is.sci.gsfc.nasa.gov/ArchivedAncillary/temporal/global/gdas
ncep_met_ct (optional)	NCEP Global Data Assimilation System (GDAS1) GRIdded Binary (GRIB) File (6 hourly, 1 degree, global). GDAS files are available in both grib1 and grib2 formats.	Current Data: ftp://is.sci.gsfc.nasa.gov/ancillary/temporal/global/gdas Archived Data: ftp://is.sci.gsfc.nasa.gov/ArchivedAncillary/temporal/global/gdas
noaa_oisst (optional)	National Oceanic and Atmospheric Administration (NOAA) Optimum Interpolation Sea Surface Temperature (OISST) (1 degree, global, weekly)	Current Data: ftp://is.sci.gsfc.nasa.gov/ancillary/temporal/global/sst Archived Data: ftp://is.sci.gsfc.nasa.gov/ArchivedAncillary/temporal/global/sst
noaa_toast (optional)	NCEP Total Ozone Analysis using SBUV/2 and TOV (TOAST) (daily, global) and TIROS Operational Vertical Sounder (TOVS) Ozone	Current Data: ftp://is.sci.gsfc.nasa.gov/ancillary/temporal/global/toast Archived Data: ftp://is.sci.gsfc.nasa.gov/ArchivedAncillary/temporal/global/toast

Input File Labels	Description	Data Sources
modis_csrb (optional)	MODIS Clear Sky Radiance Bias (CSRB). Files are available for Terra and Aqua.	Current Data: ftp://is.sci.gsfc.nasa.gov/ancillary/temporal/global/csrb Archived Data: ftp://is.sci.gsfc.nasa.gov/ArchivedAncillary/temporal/global/csrb

Parameter Labels	Description
platform	'aqua' or 'terra'
scantime	The start time of the L1B swath in yyyydddhhmm format. Note that ddd refers to the day of the year. If your input L1B files follow the standard DAAC L1B file naming convention (e.g., MYD021KM.Ayyyydddhhmmss.xxxxxxx.hdf), this information can be found in the file name itself.
destp (optional)	The destriping capability is enabled by default. Destriping removes artificial stripes that appear in the MODIS L1B 1km infrared bands due to detector-to-detector variations and mirror side effects. To disable destriping, add the 'destp' parameter with the value "DestripeOff".
gribtype (optional)	The grib1 ancillary data are used by default. To use grib2 ancillary data, add the 'gribtype' parameter with the value "2".
Output File Labels	**Description**
modis.cloudmask (optional)	Cloudmask Product output HDF file
modis.aerosols (optional)	Aerosol Product output HDF file
modis.cloudtop (optional)	Cloudtop Product output HDF file
modis.atmprofile (optional)	Atmospheric Profiles output HDF file

NOTES:

1. **Selective product processing:** The IMAPP_SPA will produce only those atmospheric products whose output labels were specified on the command line. For example, specifying only modis.cloudmask and modis.aerosol labels on the command line will result in generation of only the cloudmask and aerosol products. Output file format information for each product can be found in /SPA/IMAPP/algorithm/doc.

2. **Ancillary input files:** Specification of any ancillary input file is optional. If you do not specify a particular ancillary input label, an appropriate ancillary file will be downloaded automatically for processing. Make sure that you have an Internet connection if you do not specify the ancillary files on the command line. You may also download the appropriate ancillary files from the data sources provided in the table (see above) before execution. Recommendations for choosing ancillary input files for a particular L1B granule are provided below for each ancillary label:

   - **ssmi_nise & ssmi_seaice:** The dates for the NSIDC Near-real time Ice and Snow Extent (NISE) and NCEP sea ice datasets should be as close as possible to the dates of the L1B granules. It is recommended to use an ancillary file that is within ±14 days of the granule time. The dates for the NISE and NCEP Sea Ice ancillary files are encoded in the filenames as NISE_SSMIF13_yyyymmdd.HDFEOS and eng.yymmdd respectively. The ssmi_nise and ssmi_seaice files are required for all products.

   - **noaa_oisst:** The date of the Optimum Interpolation Sea Surface Temperature (OISST) weekly ancillary file should also correspond as closely as possible to the L1B scan time. It is recommended to use an OISST ancillary file that is within ±28 days of the granule time. The date for the OISST file is encoded in the filename as oisst.yyyymmdd. The noaa_oisst file is required for all products.

   - **noaa_toast:** The dates of the ozone data should be as close as possible to the dates of the L1B granules. For processing L1B granules dated March 31, 2005 or older, TIROS Operational Vertical Sounder (TOVS) Ozone ancillary data should be used. Otherwise Total Ozone Analysis using SBUV/2 and TOV (TOAST) data files should be used. It is recommended to use TOAST or TOVS ancillary files that are within ±14 days of the granule time. The dates for the TOAST and TOVS ozone ancillary files are encoded in the filenames as TOAST16_yymmdd.GRB and yymmdd.grb respectively. The noaa_toast file is required only for the MOD04 and MOD06 products.

- **ncep_met_ct:** Global Data Assimilation System (GDAS) files are produced every 6 hours at 00, 06, 12, and 18 UTC daily. The time, date and hour of the GDAS files in grib1 or grib2 format can be found in the filename. For example, gdas1.PGrbF00.070210.18z (gdas1.PGrbF00.070210.18z.grib2 for grib2) corresponds to February 10, 2007, 1800 UTC. When choosing the GDAS ancillary file, choose one which is closer in time (±3 hours) rather than the date. For example if you have an L1B granule at 1700 UTC, the GDAS file for 1800 hours on the same day would be the best match. However, if that is not available, it would be preferable to use the GDAS file corresponding to 1800 UTC for the day before rather than the GDAS file at 1200 UTC for the same day. It is recommended to use a gdas file for ncep_met_ct that is within ±7 days of the granule time. ncep_met_ct is required only for the MOD06 product.

- **ncep_met:** Either GDAS or Global Model Forecast Fields (GFS) files may be used for this label. Try to use a GDAS file that is within ±3 hours of the L1B granule. If that file is not available (as is often the case for real-time processing), use a GFS file instead. The naming convention for grib1 gfs files is gfs.thh.yymmdd.pgrbfxx (gfs.thh.yymmdd.pgrbfxx.grib2 for grib2). (Here yymmddd and hh represent analysis time, and xx represents forecast time step). Thus a file named gfs.t12.100201.pgrbf03 corresponds to 1500 hours (12+3) UTC on February 1, 2010. If you have to choose GFS data as input, you should attempt to use a file that is within ±1.5 hours of the L1B file. If there is more than one such GFS file, use the one with the smaller forecast time step. For example, if your data time is 15 UTC, you should try to use the 3 hour forecast field from the 1200 UTC model run, instead of the 9 hour forecast field from the 0600 UTC run. If no GDAS or GFS file is available using the above logic, use a GDAS file that is closest in time but within ±7 days of the granule time. [Note ncep_met and ncep_met_ct may end up specifying the same gdas file.] The ncep_met file is required for all products.

- **modis_csrb:** The starting date for the 8-day MODIS Clear Sky Radiance Bias (CSRB) dataset should be as close as possible to 8 days prior to the date of the L1B granules. Make sure that the CSRB file you use (Terra or Aqua) matches the platform corresponding to your L1B files. The date and platform of a CSRB file is encoded in the filename as M<Y|O>DCSR_B.Ayyyyddd.005.*.hdf (Y-Aqua; O-Terra). The modis_csrb file is required only for the MOD06 product.

Execute the 'runs': The following is an example of a command line to run the IMAPP_SPA from the /home/ipopp/drl/SPA/IMAPP/testscripts subdirectory.

```
$   ../wrapper/IMAPP/run \
    modis.mxd021km ../testdata/input/MOD021KM.10060152841.hdf \
    modis.mxd02hkm ../testdata/input/MOD02HKM.10060152841.hdf \
    modis.mxd02qkm ../testdata/input/MOD02QKM.10060152841.hdf \
    modis.mxd03 ../testdata/input/MOD03.10060152841.hdf \
    ssmi_nise ../testdata/input/NISE_SSMIF13_20100301.HDFEOS \
    ssmi_seaice ../testdata/input/eng.100301 \
    ncep_met ../testdata/input/gdas1.PGrbF00.100301.18z \
    noaa_oisst ../testdata/input/oisst.20100224  \
    noaa_toast ../testdata/input/TOAST16_100301.GRB \
    platform terra \
    scantime 2010060152841 \
    modis.cloudmask ../testdata/output/mod35.10060152841.hdf \
    modis.aerosols ../testdata/output/mod04.10060152841.hdf
```

NOTE: The above command line will produce only the cloudmask and the aerosol products.

A successful execution of 'run' usually takes some time (between 20 and 30 minutes, depending on the speed of your computer), so if the execution seems to get stuck with an "Awaiting Completion" message, do not become impatient. If execution fails, you will see an error message indicating the cause of failure (e.g., a file cannot be found, or a label cannot be recognized). Correct it and run again. If the problem has some other cause, it can be identified using the log files. Log files are automatically generated within the directory used for execution. They start with stdfile* and errfile* and can be deleted after execution. Please report any errors that cannot be fixed to the DRL. The 'run' can be executed from any directory the user chooses. This can be done by prefixing it with the file path for the 'run' script.

To Use the Script in the Testscripts Directory

One simple way to run the algorithms from any directory of your choice using your own data is to copy the test scripts from the testscripts directory to the selected directory. Change the values of the variables WRAPPERHOME, L1HOME, ANCHOME and OUTPUTDIR to reflect the file paths of the wrapper directories and the input/output directory locations. Then modify the input/output file name variables. Run the scripts to process your data.